

# Análisis de expansión de redes de telefonía móvil basándose en indicadores claves de desempeño, utilizando Procesos Gaussianos

Jhouben Janyk Cuesta Ramírez.

Director: PhD. Mauricio Alexander Álvarez.  
Facultad de Ingeniería Eléctrica, Universidad Tecnológica de Pereira

Mayo, 2013



# Contenido

## 1 Introducción

- Motivación
- Objetivos

## 2 Materiales y Métodos

- Materiales
- Métodos

## 3 Resultados

- Inferencia sobre Cada celda
- Regresión completa de la zona

# Contenido

## 1 Introducción

- Motivación

- Objetivos

## 2 Materiales y Métodos

- Materiales

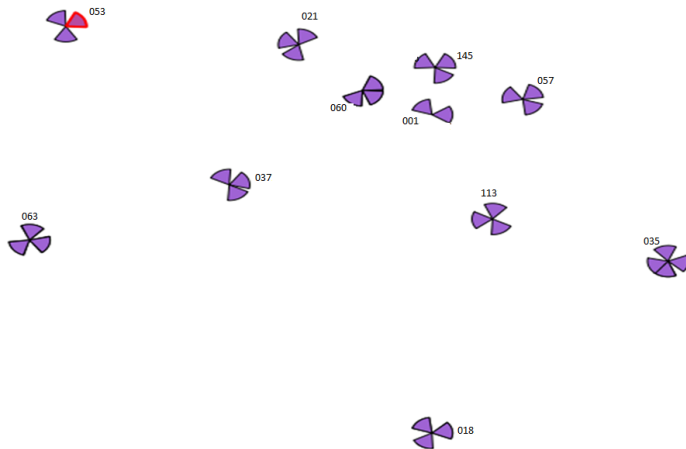
- Métodos

## 3 Resultados

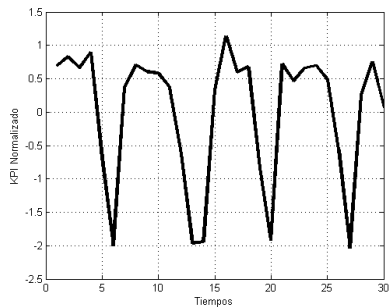
- Inferencia sobre Cada celda

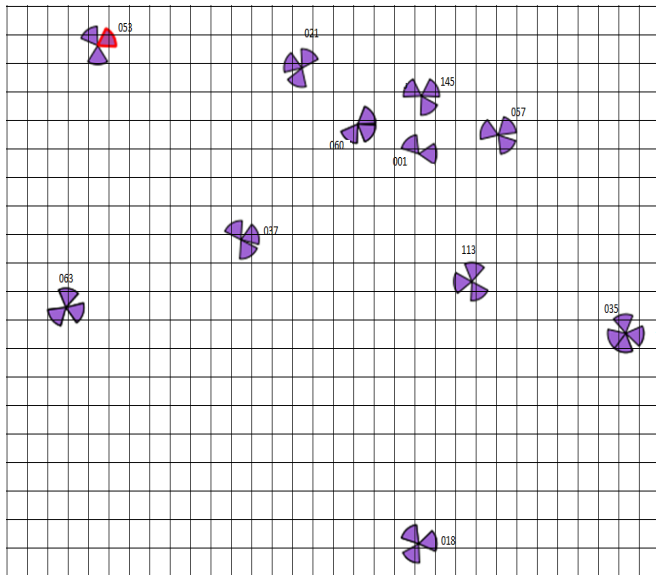
- Regresión completa de la zona

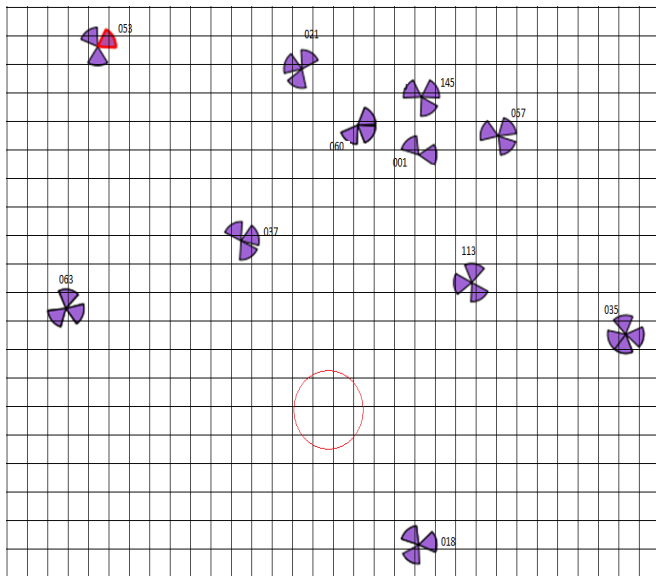
# Motivación













# Contenido

## 1 Introducción

- Motivación

- Objetivos

## 2 Materiales y Métodos

- Materiales

- Métodos

## 3 Resultados

- Inferencia sobre Cada celda

- Regresión completa de la zona

## Objetivo General

Inferir de manera acertada e inteligente, el comportamiento de un *KPI* en nuevos puntos geográficos.

## Objetivos Específicos

- Modelar un *KPI* a través de un Proceso Gaussiano espacio-temporal.
- Estimar los parámetros del modelo, usando máxima verosimilitud.
- Validar el desempeño del Proceso Gaussiano utilizando validación-cruzada.

## Objetivos Específicos

- Modelar un *KPI* a través de un Proceso Gaussiano espacio-temporal.
- Estimar los parámetros del modelo, usando máxima verosimilitud.
- Validar el desempeño del Proceso Gaussiano utilizando validación-cruzada.

## Objetivos Específicos

- Modelar un *KPI* a través de un Proceso Gaussiano espacio-temporal.
- Estimar los parámetros del modelo, usando máxima verosimilitud.
- Validar el desempeño del Proceso Gaussiano utilizando validación-cruzada.

## Objetivos Específicos

- Modelar un *KPI* a través de un Proceso Gaussiano espacio-temporal.
- Estimar los parámetros del modelo, usando máxima verosimilitud.
- Validar el desempeño del Proceso Gaussiano utilizando validación-cruzada.

# Contenido

## 1 Introducción

- Motivación
- Objetivos

## 2 Materiales y Métodos

- Materiales
- Métodos

## 3 Resultados

- Inferencia sobre Cada celda
- Regresión completa de la zona

# Materiales

- Matlab 2012a
- Microsoft Excel
- AutoCAD
- GPML toolbox.
- Windows Moviemaker



# Contenido

## 1 Introducción

- Motivación
- Objetivos

## 2 Materiales y Métodos

- Materiales
- Métodos

## 3 Resultados

- Inferencia sobre Cada celda
- Regresión completa de la zona

## El Problema de Regresión - I

El aprendizaje supervisado en regresión utiliza ejemplos

$\mathcal{D} = \{(\mathbf{x}_i, y_i) | i = 1, \dots, n\}$  para inferir nuevas salidas  $y_i = f(\mathbf{x}_i)$ .

## El Problema de Regresión - I

El aprendizaje supervisado en regresión utiliza ejemplos

$\mathcal{D} = \{(\mathbf{x}_i, y_i) | i = 1, \dots, n\}$  para inferir nuevas salidas  $y_i = f(\mathbf{x}_i)$ .

En el caso de Regresión Lineal

$$y(\mathbf{x}, \mathbf{w}) = w_0 + w_1x_1 + \dots + w_Dx_D, \quad (1)$$

## El Problema de Regresión - I

El aprendizaje supervisado en regresión utiliza ejemplos

$\mathcal{D} = \{(\mathbf{x}_i, y_i) | i = 1, \dots, n\}$  para inferir nuevas salidas  $y_i = f(\mathbf{x}_i)$ .

En el caso de Regresión Lineal

$$y(\mathbf{x}, \mathbf{w}) = w_0 + w_1x_1 + \dots + w_Dx_D, \quad (1)$$

donde  $\mathbf{x} = (x_1, \dots, x_D)^T$

## El Problema de Regresión - II

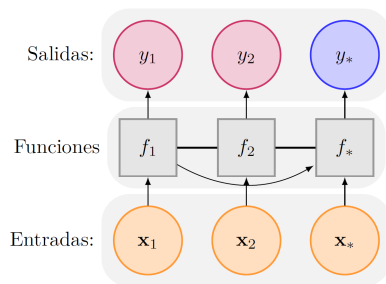
$$y(\mathbf{x}, \mathbf{w}) = \sum_{j=0}^{M-1} w_j \Phi_j(\mathbf{x}) = \mathbf{w}^T \Phi(\mathbf{x}), \quad (2)$$

donde  $\mathbf{w} = (w_0, \dots, w_{M-1})$  y  $\Phi = (\Phi_0, \dots, \Phi_{M-1})^T$  con  $\Phi_0 = 1$  y  $M$  el número de parámetros del modelo.

# Procesos Gaussianos I

# Procesos Gaussianos I

Un *Proceso Gaussiano* (GP) está compuesto por una colección de variables aleatorias escalares tal que para cada combinación finita  $\mathbf{X} = \{\mathbf{x}_1 \cdots \mathbf{x}_n\}$  las funciones  $\mathbf{f}(\mathbf{X}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}'))$  siguen una distribución Gaussiana [1].



Modelo Gráfico

## Procesos Gaussianos II

Un *GP* queda formalmente definido con la existencia de una función media y una función de covarianza

$$m(\mathbf{x}) = \mathbb{E}[f(\mathbf{x})],$$

$$k_f(\mathbf{x}, \mathbf{x}') = \mathbb{E}[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))^T],$$

El prior es de la forma  $\mathcal{N}(\mathbf{0}, K)$



## Procesos Gaussianos II

Un *GP* queda formalmente definido con la existencia de una función media y una función de covarianza

$$m(\mathbf{x}) = \mathbb{E}[f(\mathbf{x})],$$

$$k_f(\mathbf{x}, \mathbf{x}') = \mathbb{E}[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))^T],$$

El prior es de la forma  $\mathcal{N}(\mathbf{0}, K)$

## Procesos Gaussianos II

Un *GP* queda formalmente definido con la existencia de una función media y una función de covarianza

$$m(\mathbf{x}) = \mathbb{E}[f(\mathbf{x})],$$

$$k_f(\mathbf{x}, \mathbf{x}') = \mathbb{E}[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))^T],$$

El prior es de la forma  $\mathcal{N}(\mathbf{0}, K)$

## Procesos Gaussianos II

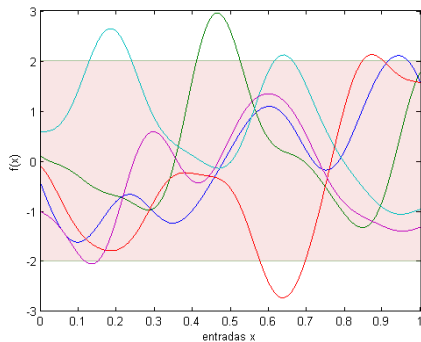
Un *GP* queda formalmente definido con la existencia de una función media y una función de covarianza

$$m(\mathbf{x}) = \mathbb{E}[f(\mathbf{x})],$$

$$k_f(\mathbf{x}, \mathbf{x}') = \mathbb{E}[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))^T],$$

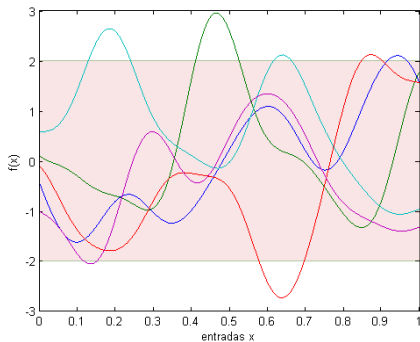
El prior es de la forma  $\mathcal{N}(\mathbf{0}, K)$

## Muestras del Prior y Posterior

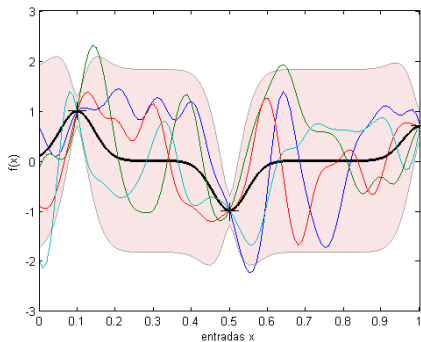


Realizaciones del Prior

## Muestras del Prior y Posterior



Realizaciones del Prior



Realizaciones del Posterior

## Predicciones I

Para observaciones ruidosas se tiene entonces el modelo

$y = f(\mathbf{x}) + \varepsilon$  , donde si el ruido es i.i.d. y con distribución de probabilidad  $\varepsilon \sim \mathcal{N}(0, \sigma_n^2)$  , la distribución conjunta de las salidas y la predicción es

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{f}_* \end{bmatrix} \sim \mathcal{N} \left( 0, \begin{bmatrix} K(X, X) + \sigma_n^2 I & K(X, X_*) \\ K(X_*, X) & K(X_*, X_*) \end{bmatrix} \right).$$

## Predicciones I

Para observaciones ruidosas se tiene entonces el modelo

$y = f(\mathbf{x}) + \varepsilon$  , donde si el ruido es i.i.d. y con distribución de probabilidad  $\varepsilon \sim \mathcal{N}(0, \sigma_n^2)$  , la distribución conjunta de las salidas y la predicción es

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{f}_* \end{bmatrix} \sim \mathcal{N} \left( 0, \begin{bmatrix} K(X, X) + \sigma_n^2 I & K(X, X_*) \\ K(X_*, X) & K(X_*, X_*) \end{bmatrix} \right).$$

## Predicciones II

La distribución condicional es de la forma

$$\mathbf{f}_* | X, \mathbf{y}, X_* \sim \mathcal{N}(\bar{\mathbf{f}}_*, \text{cov}(\mathbf{f}_*))$$

$$\bar{\mathbf{f}}_* \triangleq \mathbb{E}[\mathbf{f}_* | X, \mathbf{y}, X_*] = K(X_*, X)[K(X, X) + \sigma_n^2 I]^{-1} \mathbf{y},$$

$$\text{cov}(\mathbf{f}_*) = K(X_*, X_*) - K(X_*, X)[K(X, X) + \sigma_n^2 I]^{-1} K(X, X_*).$$



## Predicciones II

La distribución condicional es de la forma

$$\mathbf{f}_*|X, \mathbf{y}, X_* \sim \mathcal{N}(\bar{\mathbf{f}}_*, cov(\mathbf{f}_*))$$

$$\bar{\mathbf{f}}_* \triangleq \mathbb{E}[\mathbf{f}_*|X, \mathbf{y}, X_*] = K(X_*, X)[K(X, X) + \sigma_n^2 I]^{-1} \mathbf{y},$$

$$cov(\mathbf{f}_*) = K(X_*, X_*) - K(X_*, X)[K(X, X) + \sigma_n^2 I]^{-1} K(X, X_*).$$

## Funciones de Covarianza

- Función Exponencial Cuadrática [2]
- Funciones de la Clase Matérn [2]

## Función de Verosimilitud

Representa la probabilidad de las observaciones dado el modelo.

## Función de Verosimilitud

Representa la probabilidad de las observaciones dado el modelo.

$$p(\mathbf{y}|\mathbf{f}, \mathbf{X}) = \prod_{\forall i} \mathcal{N}(y_i|f_i, \sigma_n^2) \quad (3)$$

# Estimación de los Parámetros I

## Estimación de los Parámetros I

La verosimilitud marginal es de la forma :  $\mathbf{y}|\mathbf{X} \sim \mathcal{N}(\mathbf{0}, K_y)$  ,  
donde  $K_y = K + \sigma_n^2 I$ .

## Estimación de los Parámetros I

La verosimilitud marginal es de la forma :  $\mathbf{y}|\mathbf{X} \sim \mathcal{N}(\mathbf{0}, K_y)$  ,  
donde  $K_y = K + \sigma_n^2 I$ .

Tomando el logaritmo natural

$$\log p(\mathbf{y}|\mathbf{X}) = -\frac{1}{2}\mathbf{y}^T K_y^{-1}\mathbf{y} - \frac{1}{2} \log |K_y| - \frac{n}{2} \log 2\pi, \quad (4)$$

## Estimación de los Parámetros II

Derivando la Ec.4 respecto al parámetro  $\theta_j$  se obtiene

$$\frac{\partial}{\partial \theta_j} \log p(\mathbf{y}|X, \theta) = \frac{1}{2} \text{tr} \left( (\boldsymbol{\alpha} \boldsymbol{\alpha}^T - K_y^{-1}) \frac{\partial K_y}{\partial \theta_j} \right), \quad (5)$$

donde  $\boldsymbol{\alpha} = K_y^{-1} \mathbf{y}$ .



## Medidas del desempeño del modelo - I

- Error Medio Cuadrático Normalizado (SMSE) [3].
- Pérdida Logarítmica Normalizada Promedio (MSLL) [3].
- Validación Cruzada [4].

## Metodología

- Inferencia en la variable tiempo y regresión espacio-temporal para cada celda.
- Inferencia completa de la zona.

## Metodología

- Inferencia en la variable tiempo y regresión espacio-temporal para cada celda.
- Inferencia completa de la zona.

## Metodología

- Inferencia en la variable tiempo y regresión espacio-temporal para cada celda.
- Inferencia completa de la zona.

## Para tener en cuenta

- Un SMSE cercano a cero indica una buena inferencia y un MSLL negativamente alto indica un mejor modelo [5].
- Para cada validación cruzada el conjunto salidas de validación fue normalizado respecto al conjunto de entrenamiento.
- Las salidas del conjunto de entrenamiento para la inferencia completa fueron normalizadas y las coordenadas fueron re-escaladas.

## Para tener en cuenta

- Un SMSE cercano a cero indica una buena inferencia y un MSLL negativamente alto indica un mejor modelo [5].
- Para cada validación cruzada el conjunto salidas de validación fue normalizado respecto al conjunto de entrenamiento.
- Las salidas del conjunto de entrenamiento para la inferencia completa fueron normalizadas y las coordenadas fueron re-escaladas.

## Para tener en cuenta

- Un SMSE cercano a cero indica una buena inferencia y un MSLL negativamente alto indica un mejor modelo [5].
- Para cada validación cruzada el conjunto salidas de validación fue normalizado respecto al conjunto de entrenamiento.
- Las salidas del conjunto de entrenamiento para la inferencia completa fueron normalizadas y las coordenadas fueron re-escaladas.

## Espacio de Entrada - I

Espacio de entradas para regresión en  $t$  ,

$$\begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{bmatrix} \quad \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$



## Espacio de Entrada - I

Espacio de entradas para regresión en  $t$  ,

$$\begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{bmatrix} \quad \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

donde  $y_i$  equivale al valor del *KPI* en el tiempo  $t_i$ .

## Espacio de Entrada - II

Para el caso espacio temporal se tiene

$$\begin{bmatrix} x_{x,1} & x_{y,1} & t_1 \\ x_{x,1} & x_{y,1} & t_2 \\ x_{x,1} & x_{y,1} & t_n \\ \vdots & \vdots & \vdots \\ x_{x,p} & x_{y,p} & t_1 \\ x_{x,p} & x_{y,p} & t_2 \\ x_{x,p} & x_{y,p} & t_n \end{bmatrix}, \quad \begin{bmatrix} y_{1,1} \\ y_{1,2} \\ y_{1,n} \\ \vdots \\ y_{p,1} \\ y_{p,2} \\ y_{p,n} \end{bmatrix}, \quad (6)$$

## Espacio de Entrada - II

Para el caso espacio temporal se tiene

$$\begin{bmatrix} x_{x,1} & x_{y,1} & t_1 \\ x_{x,1} & x_{y,1} & t_2 \\ x_{x,1} & x_{y,1} & t_n \\ \vdots & \vdots & \vdots \\ x_{x,p} & x_{y,p} & t_1 \\ x_{x,p} & x_{y,p} & t_2 \\ x_{x,p} & x_{y,p} & t_n \end{bmatrix}, \begin{bmatrix} y_{1,1} \\ y_{1,2} \\ y_{1,n} \\ \vdots \\ y_{p,1} \\ y_{p,2} \\ y_{p,n} \end{bmatrix}, \quad (6)$$

donde,  $x_{x,i}$  y  $x_{y,i}$  indican la coordenada  $(x, y)$  de la celda  $i$  (con  $i = \{1, \dots, p\}$  con  $p = 11$  celdas), y  $t_k$  el día de la medición (con  $k = \{1, \dots, n\}$ ) siendo  $n = 30$  el número total de días).

# Contenido

## 1 Introducción

- Motivación
- Objetivos

## 2 Materiales y Métodos

- Materiales
- Métodos

## 3 Resultados

- Inferencia sobre Cada celda
- Regresión completa de la zona

## Resultados Celdas individuales

## CELDA 001

Cov.\Error	SMSE L.	SMSE G.	MSLL L.	MSLL G.
Exp C.	0.60±0.10	0.86±0.24	-0.15 ±0.21	-0.50±0.03
Màt 5	0.62±0.09	0.76±0.18	-0.20±0.11	-0.52±0.04
Màt 3	0.64±0.08	0.72±0.17	-0.20±0.09	-0.53±0.04
<b>Màt 1</b>	<b>0.69±0.05</b>	<b>0.67±0.15</b>	<b>-0.18±0.04</b>	<b>-0.54±0.05</b>

## CELDA 021

Cov.\Error	SMSE L.	SMSE G.	MSLL L.	MSLL G.
Exp C.	0.68±0.07	0.65±0.32	-0.18±0.08	-0.59±0.03
Màt 5	0.67±0.06	0.60±0.30	-0.19±0.06	-0.61±0.03
Màt 3	0.68±0.05	0.58±0.28	-0.19±0.05	-0.62±0.03
<b>Màt 1</b>	<b>0.73±0.04</b>	<b>0.55±0.25</b>	<b>-0.16±0.03</b>	<b>-0.63±0.04</b>

# Contenido

## 1 Introducción

- Motivación
- Objetivos

## 2 Materiales y Métodos

- Materiales
- Métodos

## 3 Resultados

- Inferencia sobre Cada celda
- Regresión completa de la zona

## Resultados inferencia zona geográfica completa

- Para la elección del modelo de inferencia global se eligió la función de covarianza Mètern, con parámetro  $\nu = 1$

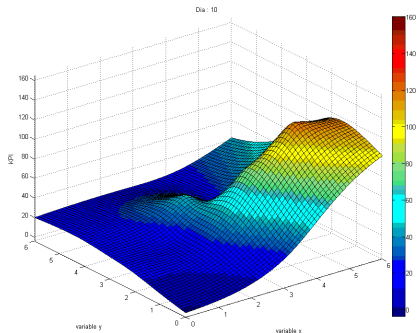
Cov.\Error	SMSE L.	SMSE G.	MSLL L.	MSLL G.
Exp C.	0.76±0.15	0.40±0.33	-0.03±0.20	-0.74±0.45
Màt 5	0.74±0.13	0.38±0.32	-0.13±0.08	-0.75±0.43
Màt 3	0.74±0.13	0.37±0.31	-0.15±0.06	-0.77±0.43
<b>Màt 1</b>	<b>0.76±0.09</b>	<b>0.35±0.29</b>	<b>-0.15±0.03</b>	<b>-0.79±0.48</b>

## Resultados inferencia zona geográfica completa

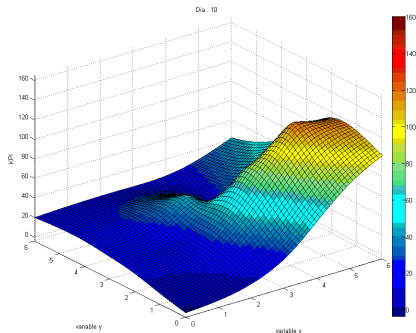
- Para la elección del modelo de inferencia global se eligió la función de covarianza Mètern, con parámetro  $\nu = 1$

Cov.\Error	SMSE L.	SMSE G.	MSLL L.	MSLL G.
Exp C.	0.76±0.15	0.40±0.33	-0.03±0.20	-0.74±0.45
Màt 5	0.74±0.13	0.38±0.32	-0.13±0.08	-0.75±0.43
Màt 3	0.74±0.13	0.37±0.31	-0.15±0.06	-0.77±0.43
<b>Màt 1</b>	<b>0.76±0.09</b>	<b>0.35±0.29</b>	<b>-0.15±0.03</b>	<b>-0.79±0.48</b>

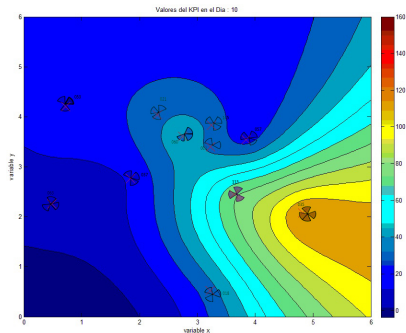




Ejemplo Superficie



Ejemplo Superficie



Ejemplo Contorno

## Conclusiones

- El modelado espacio-temporal resultó ser apropiado para los datos, aunque se presentaron pequeños valores negativos del *KPI* en algunos puntos de la zona.

## Conclusiones






- El modelado espacio-temporal resultó ser apropiado para los datos, aunque se presentaron pequeños valores negativos del *KPI* en algunos puntos de la zona.
- Se encontró una región clave, la cual obtenía los más altos valores del KPI, dicha región incluía muy pocas celdas, lo cual incentiva a pensar en ubicar una nueva celda cercana en esa zona.

## Conclusiones

- El modelado espacio-temporal resultó ser apropiado para los datos, aunque se presentaron pequeños valores negativos del *KPI* en algunos puntos de la zona.
- Se encontró una región clave, la cual obtenía los más altos valores del KPI, dicha región incluía muy pocas celdas, lo cual incentiva a pensar en ubicar una nueva celda cercana en esa zona.
- El proceso de inferencia espacio-temporal mostró como el *KPI* tiende a tener un leve comportamiento periódico.

# Gracias!

¿Preguntas?

-  L. A., B. R., and G. Richard, “Multi-dimensional signal separation with gaussian processes,” pp. 401–404, 2011.
-  K. P. Murphy, *Machine Learning a Probabilistic Perspective*. The MIT Press, 2012.
-  K. Chalupka, “Empirical evaluation of gaussian process approximation algorithms,” Master’s thesis, School of Informatics, University of Edinburgh, 2011.
-  C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
-  C. E. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning*. The MIT Press, 2006.